Why Databricks?

- **Simple** – Databricks enables a single, unified data architecture on S3 for SQL analytics, data science, and ML.
- **4x better price-performance** – Get data warehouse performance at data lake economics through SQL – optimized compute clusters.
- **Proven** – Thousands of customers have implemented Databricks on AWS to provide a game-changing analytics platform that addresses all analytics and AI use cases.
- **Trusted** – Customers in every industry trust Databricks.

Product overview

Databricks Unified Data Analytics Platform, from the original creators of Apache Spark™, unifies data science and engineering across the ML lifecycle from data preparation, to experimentation and deployment of ML applications. Databricks on AWS allows you to store and manage all your data on a simple, open lakehouse platform that combines the best of data warehouses and data lakes to unify all your analytics and AI workloads.

Healthcare Industry

Databricks enables healthcare organizations to drive innovation in the following areas.

- **Disease Prediction and Prevention** – Build powerful ML models that holistically analyze medical images, diagnostic test results, patient records, and genomic data to accurately risk score patients for chronic disease.
- **Fraud Prevention** – Build anomaly detection models on top of claims, billing, and behavioral data to identify and prevent payment of fraudulent claims, unnecessary treatments or potential identity theft.
- **Genomics and Precision Medicine** – Analyze large populations of genomic profiles to better identify disease drivers and accurate development of targeted therapeutics.
- **Hospital Operational Efficiency** – Leverage ML to drive operational efficiencies across a variety of areas such as preventing patient readmission, predicting bed utilization, etc.
- **Quality Patient Care** – Predictively determine best treatments and optimize the quality of care through the aggregation and analysis of relevant patient history and care data across all healthcare channels.

Product features

Databricks includes the following features:

- Reliable data engineering
- SQL Analytics on all your data
- Collaborative data science
- Production ML

Use Cases

Below are the most popular use cases for AWS:

- **Personalized recommendation engines** – Process your data in real time to provide the most relevant product and service recommendations.
- **Genomic sequencing** – Modernize your technology stack to improve experience for patients and physicians with the fastest DNASeq pipeline at scale.
- **Fraud detection and prevention** – Leverage complete historical data together with real-time data streams to quickly identify anomalous and suspicious financial transactions.

Databricks Unified Data Analytics Platform

Data Analytics and ML (ML) for the Healthcare Industry
How it works

Databricks provides a unified open platform for all your data. It empowers data scientists, data engineers, and data analysts with a simple collaborative environment to run interactive and scheduled data analysis workloads.

Databricks is from the original creators of some of the world’s most popular open-source projects, Apache Spark, Delta Lake, MLflow, and Koalas. It builds on these technologies to deliver a true lake house architecture combining the best of data lakes and data warehouses for a fast, scalable, and reliable data platform.

Built for the cloud, your data is stored in low-cost cloud object stores such as Amazon Simple Storage Service (Amazon S3) with performant access enabled through a caching data layout and indexing layer to work with your data you can launch clusters with hundreds of machines.

To work with your data, you can launch clusters with hundreds of machines each with a mixture of CPUs and GPUs needed for your analysis. If you're on a large data team, policies can define limits on cluster sizes and configuration. There is a Databricks runtime for Data Engineers and Data Scientists, as well as a runtime that is optimized for ML workloads.
What our customers are saying

"Databricks allows us to take clinical research and development and turn it into a clinically validated screen in far less time, which allows us to save us a lot of the money and effort it would generally require to do this on our own on-premises computational platforms."

—Lynn Carmichael, Senior Director of Computational Bioinformatics, Sanford Health

Additional Resources

1. Databricks on AWS
2. Sanford Health Case Study
3. Product overview

Solution available in AWS Marketplace